Available online at www.sciencedirect.com

ScienceDirect

Intelligence 34 (2006) 479–486

INTELLIGENCE

Males have greater *g*: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test[☆]

Douglas N. Jackson, J. Philippe Rushton *

The University of Western Ontario, Canada

Received 29 August 2005; received in revised form 17 February 2006; accepted 11 March 2006
Available online 21 April 2006

Abstract

In this study we found that 17- to 18-year old males averaged 3.63 IQ points higher than did their female counterparts on the 1991 Scholastic Assessment Test (SAT). We analysed 145 item responses from 46,509 males and 56,007 females (total $N = 102{,}516$) using a principal components procedure. We found (1) the *g* factor underlies both the SAT Verbal (SAT-V) and the SAT Mathematics (SAT-M) scales with the congruence between these components greater than 0.90; (2) the *g* components predict undergraduate grades better than do the traditionally used SAT-V and SAT-M scales; (3) the male and the female *g* factors are congruent in excess of .99; (4) male–female differences in *g* have a point-biserial effect size of 0.12 favoring males (equivalent to 3.63 IQ points); (5) male–female differences in *g* are present throughout the entire distribution of scores; (6) male–female differences in *g* are found at every socioeconomic level; and (7) male–female differences in *g* are found across several ethnic groups. We conclude that while the magnitude of the male–female difference in *g* is not large, it is real and non-trivial. Finally, we discuss some remaining sex-difference/brain-size/IQ anomalies.

Keywords: Brain size; Intelligence tests; Gender differences; Evolutionary psychology; Psychometrics

1. Introduction

For approximately a century a consensus has existed that there are no sex differences in overall general intelligence. It was British psychologist Cyril Burt who first advanced this conclusion based on the results from a series of reasoning tasks he developed and administered to both boys and girls in various secondary schools in Liverpool (inspired partly by Galton's work and that of Binet in France; Burt & Moore, 1912). He thereby overturned traditional Victorian wisdom (Ellis, 1904). Terman (1916, pp. 69–70) further advanced the conclusion with his American standardization sample of the Stanford–Binet on approximately 1000 4- to 16-year-olds. Their findings of "no sex difference in intelligence" have since been replicated many times on other standardization samples with other test batteries. However, males are often observed to average higher scores on some tests of spatial ability, mathematical reasoning, and targeting, while females are often found to average higher on some tests of memory, verbal ability, and motor coordination within personal space (Halpern, 2000; Kimura, 1999). Also, males have been found to have

---

[☆] D.N.J. carried out the statistical analyses of these data and presented them at the December 2002 meeting of the International Society for Intelligence Research (ISIR) in Nashville, TN. Following D.N.J.'s death in September 2004, J.P.R. completed the write up as presented in this article.

* Corresponding author. Department of Psychology, University of Western Ontario, London, Ontario Canada N6A 5C2. Tel.: +1 519 661 3685; fax: +1 519 850 2302.

E-mail address: Rushton@uwo.ca (J.P. Rushton).

0160-2896/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2006.03.005

greater test score variance than females, being overrepresented at both the high and the low extremes (Deary, Thorpe, Wilson, Starr, & Whalley, 2003; Hedges & Nowell, 1995). When marked sex differences are found in particular abilities, such as the 1 standard deviation male advantage in rotating imaginary objects (Kimura, 1999; Voyer, Voyer, & Bryden, 1995), they are typically considered "sex-biased" and excluded from general test batteries.

Two recent sets of observations raise anew the question of sex differences in general intelligence in normal populations. The first is that the general factor of mental ability-*g*-permeates all tests to a greater or lesser extent, as initially proposed by the British psychologist Charles E. Spearman (1904, 1923; Jensen, 1998). More than any other factor, the magnitude of the test's *g* loading best determines a test's power to predict academic achievement, creativity, career potential, and job performance (Kuncel, Hezlett, & Ones, 2004; Lubinski, 2004). Thus, a "spatial" test may be relatively high on *g* (mental rotation) or low (perceptual speed), a "verbal" test may be relatively high (reasoning) or low (fluency), as may a "memory" test be high (repeating a series in reverse order) or low (repeating a series in presented order). As Spearman (1923, p. 198) noted, there is an "indifference of the indicator," meaning that the specific content or form of the item is inconsequential so long as it is apprehended or perceived in the same way by all persons taking the test. The question of sex differences in general intelligence can therefore be formulated more precisely as: "Are there sex differences on the *g* factor?"

The second set of observations concerns the sex difference found in brain size and the relation between brain size and cognitive ability. In 1992, C. Davison Ankney re-analysed 1000 brain weights at autopsy from a study published by Ho, Roessmann, Straumfjord and Monroe (1980) and discovered that males averaged a larger brain size than females even after adjusting for body size (140 g before adjustments; 100 g after adjustments). Ankney's findings were immediately corroborated by Rushton (1992) using cranial capacity data calculated from head size measurements gathered from a stratified random sample of 6325 U.S. military personnel with individual adjustments made for many different body size variables using analysis of covariance. At the time, Ankney's and Rushton's findings were considered "revolutionary" (Maddox, 1992) because prior studies of sex differences in brain size argued that they "disappeared" when adjusted for body size (by using an inappropriate adjustment based on brain- to body-size ratios, it turns out; Ankney, 1992). Additional data on sex differences in brain size were reviewed by Rushton and Ankney (1996) who confirmed the male advantage; they also reviewed the relation between brain size and cognitive ability, which they found to be 0.20 using external head size measures and 0.40 using magnetic resonance imaging, or MRI. Subsequently, Pakkenberg and Gundersen (1997) documented that men have 15% more neurons than women (22.8 versus 19.3 billion), and over two-dozen MRI studies have confirmed the brain-size/IQ correlation of about 0.40 (e.g., Ivanovic et al., 2004; McDaniel, 2005).

Lynn (1999, p. 1) dubbed the findings on sex differences in brain size "the Ankney–Rushton anomaly." He argued that if brain size is linked to IQ, and males average larger brains than females, then men should have higher average intelligence than women. Lynn (1994, 1999) reviewed data from a number of published tests such as the well-standardized highly-*g*-loaded Wechsler Adult Intelligence Scale-Revised (WAIS-R) from countries as varied as Britain, Belgium, Greece, China, Israel, the Netherlands, Norway, Sweden, Japan, India, and Indonesia, as well as the United States, and found that men average 3.8 IQ points higher than women. Also in support, Lynn and Irwing (2004) carried out a meta-analysis of 57 studies of general population samples to examine sex differences on the Standard and Advanced Progressive Matrices, one of the most highly-*g*-loaded tests of non-verbal reasoning, and found that adult men exceeded adult women by an average of 5.0 IQ points. Subsequently, Irwing and Lynn (2005) carried out a meta-analysis of 22 studies of university samples on the Progressive Matrices and found the male advantage averaged between 3.3 and 5.0 IQ points, with 4.6 being their best estimate.

Lynn (2005) has also pointed to other research supportive of his hypothesis, such as that by Baron-Cohen (2003; who prefers the terminology that men have greater "synthesising ability"). He cited studies showing that people of both sexes consistently rate their fathers as more intelligent than their mothers (Furnham, 2001) and of self-ratings showing that males rate themselves higher than do females even after adjusting for any effects of sex differences in personality (Furnham & Buchanan, 2005). He also pointed to the evidence from history that males made 98% of the world's contributions to knowledge (Murray, 2003).

Age turns out to be an important factor for determining sex differences in IQ because the male advantage does not emerge until the late adolescent growth spurt when brain size differences peak. Girls mature faster than boys, which give them an early advantage in language development and may mask later cognitive differences. One study of 1400 3-year-olds in Mauritius found girls had a nearly

2 IQ point advantage over boys in verbal ability and also in visual–spatial ability on the Boehme Test of Basic Concepts (Lynn, Raine, Venables, & Mednick, 2005). However, among children 5 to 15 years of age, Lynn (1994, 1999) found no consistent sex differences, a general finding which he suggested led generations of researchers, who relied on school samples, to miss the later emerging sex difference.

Lynn's procedure for determining sex differences in *g* has been the straightforward one of summing items and subtests from *g* loaded batteries to produce a total score. Jensen (1998) criticized this method as producing outcomes dependent on the particular mix of questions in the battery and suggested the use of principal components analysis to extract *g* instead, a procedure he had found useful in examining ethnic and other group differences. When Jensen (1998) carried out this analysis with five large data sets he observed considerable inconsistency: although males averaged higher in three studies (by 5.49, 2.81, and 0.18 IQ points), females averaged higher in two others (by 7.91 and 0.11 IQ points), which he averaged to give a negligible male advantage of 0.11 IQ points.

Jackson (2002) endorsed Jensen's use of the principal components procedure but criticized his choice of tests, suggesting that the General Aptitude Test Battery (the one on which females had scored 7.91 IQ points higher), biased the results because it included a large number of low *g* tests of mechanical aptitude on which males excelled but which resulted in their achieving a low *g* score. He suggested limiting the use of principal components analysis to highly-*g*-loaded tests such as the Wechsler or his own Multi-Dimensional Aptitude Battery (Jackson, 1984). When Jackson (2002) analysed previously published data this way, including that by Jensen and Reynolds (1982) on the U.S. standardization of the Wechsler Intelligence Scale for Children-Revised (WISC-R), and by Lynn and Mulhern (1991) on the Scottish WISC-R standardization, he found that males outperformed females on the more *g* loaded subtests.

Most recently, Nyborg (2003, 2005) has argued that the principal components type of analysis used by Jensen and Jackson is insufficiently sensitive to detect true score differences and so provides little or no improvement over Lynn's summing of subtests because the magnitude of the sex difference is modest and the number of subtests being analysed is typically 12 or fewer. In this type of analysis the test of significance is often based on the number of subtests and so the result can be thrown off by the inclusion or exclusion of even one subtest. Nyborg noted that although more powerful multivariate methods (such as hierarchical factor analysis and structural equation modeling) are available for identifying and controlling factorial impurities in the measures employed (e.g., spatial visualization rather than *g*) and although these are used routinely when looking at other group differences, they are rarely used in studying sex differences.

To date, only two studies have been entirely satisfactory from a multivariate perspective and both have other imperfections. Nyborg (2005) used hierarchical factor analysis with the Schmid–Leiman transformation on a battery of 20 widely differing ability tests given to carefully matched groups of Danish 17-year-olds and found a *g* difference favoring males equivalent to 8.55 IQ points. Although this was a longitudinal study with repeated measures, the main analyses rested on a very small sample of only 31 males and 31 females. The other study, by Colom, Garcia, Juan-Espinosa and Abad (2002), analysed 703 female and 666 male participants in the Spanish standardization of the WAIS-III and reported a male advantage of 3.6 IQ points. The clarity of the result was reduced, however, when the authors combined it with other findings and concluded there were "null sex differences in general intelligence." It was only when Nyborg (2003, 2005) disaggregated the data of Colom et al. (2002) and applied a statistical test that the male advantage on *g* emerged.

2. Method

The present study provides a more definitive test of the hypothesis that males average higher in *g* than females by factor analysing the 145 item scores (correct or incorrect) of 46,509 male and 56,007 female 17- to 18-year-olds (total $N = 102{,}516$) from the "validity study sample" of the 1991 administration of the well-known Scholastic Assessment Test (SAT; College Entrance Examination Board, 1992). Before 1990 the test was called the Scholastic Aptitude Test; after 1994 it was renamed simply SAT—not an acronym.

The SAT provides a reasonable way of testing the hypothesis of sex differences in *g*. It is an exam that several generations of high school students have taken for admission to college in the United States that was carefully developed from a psychometric perspective to maximize prediction of college-level academic achievement while minimizing extraneous factors such as potential sex, ethnic, and social class bias. About 50% of the U.S. population now go to college and SAT test takers are representative of those who aspire to do so. Over decades of research it has demonstrated substantial reliability and validity and is considered the "gold standard" for academic achievement tests (The lowest scoring SAT group are likely below the average IQ of 100.).

Although the SAT is widely believed to measure mainly academic achievement, Frey and Detterman (2004) have shown that it is an excellent measure of general cognitive ability (*g*). For example, they found that scores on the SAT correlated highly (0.82) with *g* extracted from the Armed Services Vocational Aptitude Battery (ASVAB), as it also did with *g* loaded mental ability tests such as the California Test of Mental Maturity (0.82), the Otis–Lennon Mental Ability Test (0.78), the Lorge–Thorndike Intelligence Test (0.79), the Differential Aptitude Test (0.78), the Henmon–Nelson Test of Mental Maturity (0.65), the Coop School and College Ability Test (0.53), and the Raven's Advanced Progressive Matrices (0.48; 0.72 when corrected for restricted range).

The item scores to be analysed are the responses to multiple-choice questions arranged in four sections, each lasting 30 min: a 45-item verbal section, a 40-item verbal section, a 25-item mathematics section, and a 35-item mathematics section. The verbal questions were of four types: analogies, reading comprehension, antonyms, and sentence completions; the mathematics questions were also of four types: arithmetic, algebra, geometry, and quantitative comparisons. These sum to yield two scores: SAT Verbal (SAT-V) and SAT Mathematics (SAT-M). However, we disregard question type, section, and subtest in most of our analyses and instead focus on extracting the *g* component from each item. Such procedures have proven very useful in establishing the convergent and divergent validities of mental ability subtests as well as their *g* loadings (Jackson, 1984; Jensen, 1998). They have also proven useful for determining the structure of intelligence in different ethnic groups (Jensen, 1998; Rushton & Jensen, 2005; Rushton, Skuy, & Bons, 2004).

3. Results

Both the SAT Verbal (SAT-V) scale and the SAT Mathematics (SAT-M) scale have a mean of 500 and a standard deviation of 100. In our 1991 validity sample, males averaged 499 on the SAT-M and 434 on the SAT-V, while females scored 457 on the SAT-M and 425 on the SAT-V—a magnitude of sex difference reported for the last 32 years (Halpern, 2000, p. 129, Figs. 3.8 and 3.9). In 1995, both the SAT-M and SAT-V scores were re-centered upwards by a total of about 100 points or one standard deviation to adjust for the fact that the mean scores had been declining.

We first carried out a principal components analysis of the items for the SAT-V and SAT-M subtests to extract their corresponding *g* scores. Table 1 provides the correlations between scores from the SAT-V and the SAT-M and their corresponding first principal components. Note that while the correlation between SAT-V and SAT-M is 0.67, the corresponding correlation for their *g* scores is 0.90. Given their respective reliabilities, this latter value approaches unity when corrected for attenuation. These results support Frey and Detterman's (2004) conclusion that the SAT-V and SAT-M are primarily measures of *g*, with the other factors they measure being secondary. When sex is entered into the matrix as a dummy variable, the point-biserial correlations are positive, indicating that males score higher on average than females. Extracting the *g* factor score from all the SAT-V and SAT-M items together gave a point-biserial effect size favoring males of 0.12, which is equivalent to 0.24 standard deviation units (see Jensen, 1998, p. 543, n. 12). Given a standard deviation of 15 typical for tests such as the Wechsler, the difference between men and women is equivalent to 3.63 IQ points.

Table 1
Correlations between sex*, SAT scale scores, and SAT *g* scores (*N*=102,516)

	Sex	SAT-V	SAT-M	SAT-V *g* score	SAT-M *g* score
Sex	1.00	0.05	0.19	0.10	0.12
SAT-V		1.00	0.67	0.91	0.89
SAT-M			1.00	0.86	0.84
SAT-V *g* score				1.00	0.90
SAT-M *g* score					1.00

Note: *Females = 1, males = 2.

The *g* factors extracted from the male and female samples were highly congruent, in fact virtually identical, in excess of 0.99. This means that the items "behave" in the exact same way for males as they do for females and have the same "meaning." For example, those items that men find difficult, women do too.

We examined how well the SAT-V and SAT-M scores predict freshman year university grades before and after *g* has been partialed out. Fig. 1 depicts the correlations between SAT-V and SAT-M scores and grades for English, Foreign Language, Natural Science, Mathematics, and Social Science. The first two bars in each set are those for the traditional SAT-V and SAT-M scales. The second two bars are the respective correlations with the *g* factor removed. Note that the non-*g* components make a negligible contribution for all academic subjects except Mathematics. For Mathematics there is a modest residual correlation of approximately 0.12, suggesting that, in addition to *g*, differences in experience with mathematics or some residual component such as spatial ability might contribute to correlations with grades.

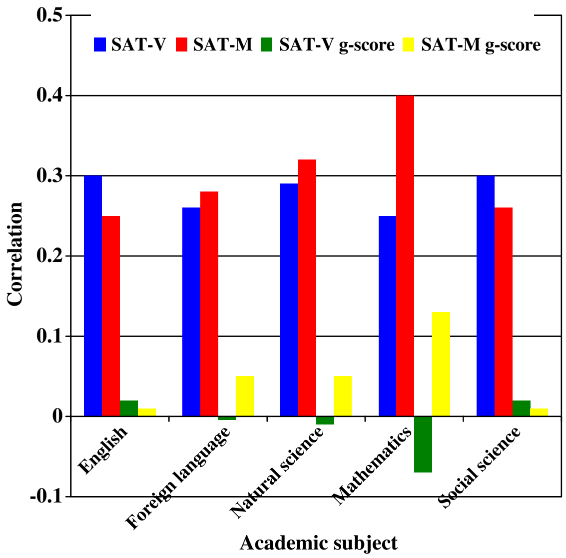

Fig. 1. SAT correlations with grades. The effects of removing *g* factor variance from SAT scores reduces correlation to zero.

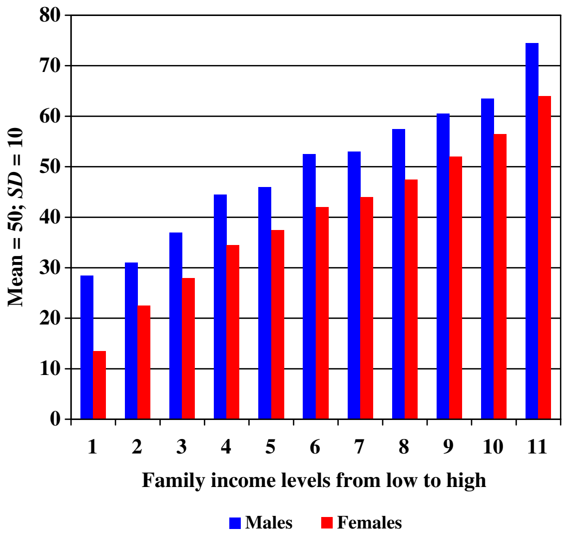

Fig. 3. Male and female SAT *g* scores (mean standardized) for eleven levels of family income.

We examined the male–female differences in *g* across a representation of the entire distribution (Fig. 2). Males are over-represented in each and every category above the middle category, and females are over-represented in each and every category below the middle category. This indicates that male–female differences in *g* occur across the entire distribution of *g* scores. Category 1, representing the lowest block of scores, contains persons who scored below chance (perhaps misled by distracter items). The ratio of females to males who did so was approximately 5:3. In the past, when males were observed to obtain a higher mean score, it was often attributed to males having greater representation at the extreme high end of the distribution (Hedges & Nowell, 1995). Fig. 2 indicates this is not correct.

Fig. 3 contains the male and female *g* score distributions (mean standardized to 50, with SD of 10) for eleven levels of family income. Note that the male–

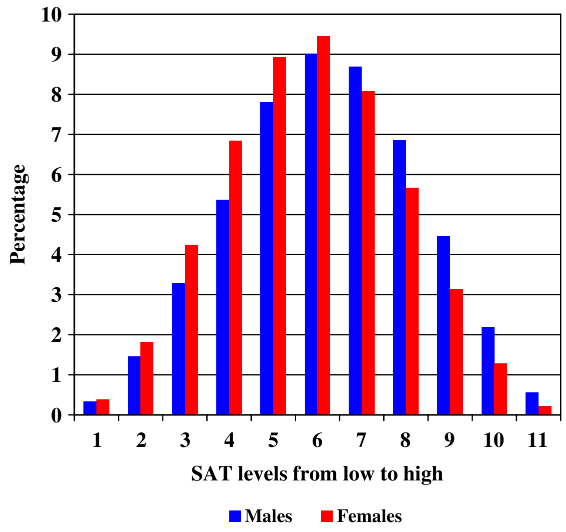

Fig. 2. Distribution of males and females at eleven SAT *g* factor score levels. Males are more numerous at upper range of the distribution and females in the lower range.

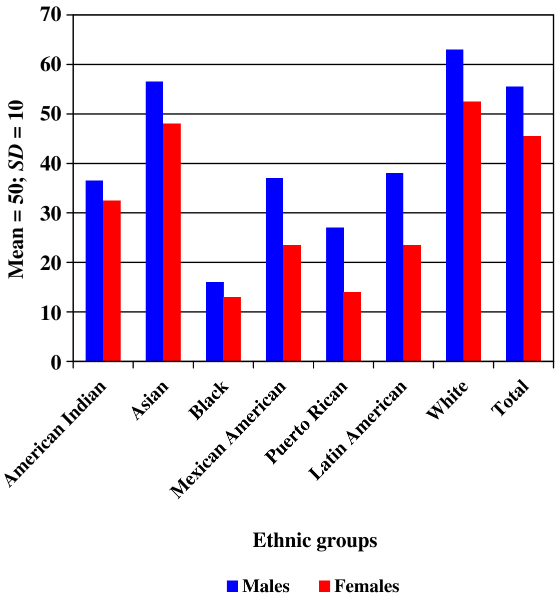

Fig. 4. Male and female SAT *g* scores (mean standardized) for seven different ethnic groups.

female difference holds at each level. The same conclusion of higher scores for males also holds for nine levels of father's education and nine levels of mother's education (not shown as figures). Hence, it is most unlikely that the sex differences are due to these type of family background factors.

Fig. 4 shows the overall male–female *g* score differences (mean standardized to 50, with SD of 10) across seven ethnic groups, representing the largest broad groups who complete the SAT. All seven groups show sex differences although there is variation among them on the magnitude of the difference. The two far right bars show the overall mean is 105 for men and 95 for women.

4. Discussion

This study found a point-biserial effect size of 0.12 favoring males on the SAT, which provides a good measure of *g* as manifested through "school-learned abilities" in high school graduating samples. (This is equivalent to a *d* of 0.24 on a test with a conventional SD of 15, or 3.63 IQ points.) Despite its modest size, particularly in comparison with the variance of scores within the sexes, the results reported here support Lynn's (1994, 1999) hypothesis of significant sex differences in average cognitive ability. It is in accord with the previous estimates of the male advantage reviewed in the Introduction and approximates the 3.8 IQ points estimated by Lynn (1999) but shy of the 5.0 IQ points estimated by Lynn and Irwing (2004) using the Progressive Matrices. However, Lynn and Irwing's analyses gave a 0.16*d* for 17-year-olds (IQ = 2.4) and a 0.30*d* for 20- to 29-year-olds (IQ = 4.5) so our result for 17-year-olds corroborates Lynn's rather closely for this age range.

A point-biserial effect size of 0.12 is considered "small." Nonetheless, it indicates that if selection occurred at the mean, then 55% of males would pass compared to only 45% of females (Rosnow, Rosenthal, & Rubin, 2000). If a more stringent criterion for selection were applied, say at the 85th percentile, the ratio of males selected over females would be considerably higher. Moreover, Jackson (2002) suggested that because test constructors such as himself and the Educational Testing Service (which developed the SAT) often eliminate items showing marked sex differences in order to reduce the perception of bias, it is possible that the results reported here might be a lower-bound estimate of the "true" sex difference. It is also possible that the male advantage is underestimated because of restriction in range on *g* in this higher performance group; if so, the sex difference would be larger in the general population.

It might be questioned, however, whether findings about sex differences on the high-end SAT are generalizable to the general population. Also, because 55% of our SAT sample was female, it might be hypothesized that we drew from lower in the female IQ pool, and that this is why a lower mean *g* was found for females. Nonetheless, as shown in Figs. 2, 3, and 4, the sex difference was found in the very highest and the very lowest SAT levels, at the very highest and the very lowest socioeconomic status levels, and for each and every ethnic group, including Whites examined alone. To maintain that selection bias caused the sex difference in this data set, therefore, would require the assumption that there are hypothetical respondents who, if tested, would provide a compensating female–male advantage in *g* that would counterbalance the findings. They would have to be found at every level of SAT performance, in every level of family income, for every level of fathers' and of mothers' education, and for every ethnic group examined.

It could still be argued, however, that some ambiguity remains in interpreting the results. Although there is a clear sex difference in *g* among students taking the test, it is perhaps less clear that this sex difference captures faithfully the sex difference in the general population. Better sampling of both sexes from the SAT pool might answer this question more definitively, as may additional parametric studies of sex differences from other parts of the *g* distribution.

The *g* factors extracted from the male and female samples were highly congruent indicating they were measuring the same thing for both sexes. Sex differences in *g* probably rest on sex differences in brain functioning. Both Ankney (1992) and Rushton (1992) found that males average 100 g more brain weight than females even after correction for body size. Ankney estimated the sex difference in brain size as 0.78 standard deviations. Assuming a 0.35 correlation between brain size and IQ, therefore, Lynn (1999) and Nyborg (2005) predicted that the male advantage in average IQ is $0.78 \times 0.35 = 0.273$ standard deviation units, equal to 4.10 IQ points on a test like the Wechsler standardized with SD = 15. This predicted outcome is very close to the 4.3 mid-point of the range of observed outcomes for large samples, i.e., IQ = 3.6 in the present study and IQ = 5.0 in Lynn and Irwing's (2004) review.

However, several sex-difference/brain-size/IQ anomalies still require resolution. First, there is a major gap in Lynn's resolution of the Ankney–Rushton anomaly. Males are found to average a larger brain size *from birth onwards, even after controlling for body size*. For example, in a study of 100 East Asian children followed from birth to age seven, boys at birth averaged a cranial capacity 5 cm^3 larger than girls, a difference that increased to 40 cm^3 by 4 months, and 50 cm^3 by age 1 year and then remained stable through to age 7 years (Rushton, 1997; controlling for body size). Other data show the 50 cm^3 male advantage in brain size at 1 year remains stable until adolescence when male brains grow to become 140 to 160 cm^3 larger than female brains (Rushton & Ankney, 1996; also controlling for body size).

Brain tissue is metabolically expensive. It would be interesting to know what these "extra male neurons" are doing from birth to age 16 before the male advantage in IQ manifests itself. Ankney (1995) hypothesized that they are related to the male advantage in dynamic spatial abilities (not measured by IQ tests) such as in throwing at targets. A male advantage at targeting shows up among 3- to 5-year-olds even when the tasks are simple underhanded throws and avoid sex differences in skeletomuscular structure (Kimura, 1999). Dynamic spatial ability may also explain the additional anomaly that Black males average a larger brain size than do White or East Asian females, even after adjustments for body size, and despite averaging 11 IQ points or more lower than White or East Asian females (Rushton & Ankney, 1996). Additional research using magnetic resonance imaging and a wider array of cognitive tasks could surely untangle these further conundrums and interrelationships.

References

Ankney, C. D. (1992). Sex differences in relative brain size: The mismeasure of woman, too? *Intelligence*, *16*, 329−336.

Ankney, C. D. (1995). Sex differences in brain size and mental abilities: Comments on R. Lynn and D. Kimura. *Personality and Individual Differences*, *18*, 423−424.

Baron-Cohen, S. (2003). *The essential difference.* London: Penguin/Perseus.

Burt, C. L., & Moore, R. C. (1912). The mental differences between the sexes. *Journal of Experimental Pedagogy*, *1*, 273–284, 355–388.

College Entrance Examination Board. (1992). *Validity study sample of the 1991 SAT administration.* New York, NY: College Entrance Examination Board.

Colom, R., Garcia, L. F., Juan-Espinosa, M., & Abad, F. (2002). Null sex differences in general intelligence: Evidence from the WAIS-III. *Spanish Journal of Psychology*, *5*, 29−35.

Deary, I. J., Thorpe, G., Wilson, V., Starr, J. M., & Whalley, L. J. (2003). Population sex differences in IQ at age 11: The Scottish Mental Survey 1932. *Intelligence*, *31*, 533−542.

Ellis, H. (1904). *Man and woman* (4th edition). New York: Houghton Mifflin.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science*, *15*, 373−378.

Furnham, A. (2001). Self-estimates of intelligence. *Personality and Individual Differences*, *31*, 1381−1405.

Furnham, A., & Buchanan, T. (2005). Personality, gender and self-perceived intelligence. *Personality and Individual Differences*, *39*, 543−555.

Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd edition). Mahwah, NJ: Erlbaum.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41−45.

Ho, K. C., Roessmann, U., Straumfjord, J. V., & Monroe, G. (1980). Analysis of brain weight: I and II. *Archives of Pathology and Laboratory Medicine*, *104*, 635−645.

Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the Progressive Matrices in university students: A meta-analysis. *British Journal of Psychology*, *96*, 505−524.

Ivanovic, D. M., Leiva, B. P., Castro, C. G., Olivares, M. G., Jansana, J. M. M., Castro, V. G., et al. (2004). Brain development parameters and intelligence in Chilean high school graduates. *Intelligence*, *32*, 461−479.

Jackson, D. N. (1984). *Multidimensional aptitude battery.* Port Huron, MI: Research Psychologists Press.

Jackson, D. N. (2002, December 5–7). *Evaluating g in the SAT: Implications for the sex differences and interpretations of verbal and quantitative aptitude*. Paper presented at the International Society for Intelligence Research, Nashville, TN.

Jensen, A. R. (1998). *The g factor.* Westport, CT: Praeger.

Jensen, A. R., & Reynolds, C. R. (1982). Race, social class and ability patterns on the WISC-R. *Personality and Individual Differences*, *3*, 423−438.

Kimura, D. (1999). *Sex and cognition.* Cambridge, MA: MIT Press.

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, *86*, 148−161.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman's (1904) "'General intelligence,' objectively determined and measured". *Journal of Personality and Social Psychology*, *86*, 96−111.

Lynn, R. (1994). Sex differences in intelligence and brain size: A paradox resolved. *Personality and Individual Differences*, *17*, 257−271.

Lynn, R. (1999). Sex differences in intelligence and brain size: A developmental theory. *Intelligence*, *27*, 1−12.

Lynn, R. (2005). Sex differences in IQ. *The Psychologist*, *18*, 471 (Letter).

Lynn, R., & Irwing, P. (2004). Sex differences on the Progressive Matrices: A meta-analysis. *Intelligence*, *32*, 481−498.

Lynn, R., & Mulhern, G. (1991). A comparison of sex differences on the Scottish and American standardization samples of the WISC-R. *Personality and Individual Differences*, *4*, 1179−1182.

Lynn, R., Raine, A., Venables, P. H., & Mednick, S. A. (2005). Sex differences in 3 year olds on the Boehm Test of Basic Concepts: Some data from Mauritius. *Personality and Individual Differences*, *39*, 683−688.

Maddox, J. (1992). How to publish the unpalatable? *Nature*, *358*, 187.

McDaniel, M. A. (2005). Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence*, *33*, 337−346.

Murray, C. (2003). *Human accomplishment.* New York: HarperCollins.

Nyborg, H. (2003). Sex differences in *g*. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen,* (pp. 187−222). Oxford: Elsevier.

Nyborg, H. (2005). Sex-related differences in general intelligence *g*, brain size, and social status. *Personality and Individual Differences*, *39*, 497−509.

Pakkenberg, B., & Gundersen, H. J. G. (1997). Neocortical neuron number in humans: Effects of sex and age. *Journal of Comparative Neurology*, *384*, 312−320.

Rosnow, R. L., Rosenthal, R., & Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological Science*, *11*, 446−453.

Rushton, J. P. (1992). Cranial capacity related to sex, rank, and race in a stratified random sample of 6325 U. S. military personnel. *Intelligence*, *16*, 401−413.

Rushton, J. P. (1997). Cranial size and IQ in Asian Americans from birth to age seven. *Intelligence*, *25*, 7−20.

Rushton, J. P., & Ankney, C. D. (1996). Brain size and cognitive ability: Correlations with age, sex, social class, and race. *Psychonomic Bulletin and Review*, *3*, 21−36.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law*, *11*, 235−294.

Rushton, J. P., Skuy, M., & Bons, T. A. (2004). Construct validity of Raven’s advanced progressive matrices for African and non-African engineering students in South Africa. *International Journal of Selection and Assessment*, *12*, 220−229.

Spearman, C. (1904). “General intelligence,” objectively determined and measured. *American Journal of Psychology*, *15*, 201−292.

Spearman, C. (1923). *The abilities of man.* New York: MacMillan.

Terman, L. M. (1916). *The measurement of intelligence.* Boston, MA: Houghton Mifflin.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin*, *117*, 250−270.